

SEQUENCE LISTING

<110> KYOWA HAKKO KOGYO CO., LTD

<120> Novel Transaldolase

<130> 00005.001198

<140> US/10/088,594

<141> 2002-03-21

<150> JP 99/266548

<151> 1999-09-21

<160> 3

<170> PatentIn Ver. 2.0

<210> 1

<211> 1080

<212> DNA

<213> Corynebacterium glutamicum ATCC31388

<400> 1

atg tct cac att gat gat ctt gca cag ctc ggc act tcc act tgg ctc 48
Met Ser His Ile Asp Asp Leu Ala Gln Leu Gly Thr Ser Thr Trp Leu
1 5 10 15

gac gac ctc tcc cgc gag cgc att act tcc ggc aat ctc agc cag gtt 96
Asp Asp Leu Ser Arg Glu Arg Ile Thr Ser Gly Asn Leu Ser Gln Val
20 25 30

att gag gaa aag tct gta gtc ggt gtc acc acc aac cca gct att ttc 144
Ile Glu Glu Lys Ser Val Val Gly Val Thr Thr Asn Pro Ala Ile Phe
35 40 45

gca gca gca atg tcc aag ggc gat tcc tac gac gct cag atc gca gag 192
Ala Ala Ala Met Ser Lys Gly Asp Ser Tyr Asp Ala Gln Ile Ala Glu
50 55 60

ctc aag gcc gct ggc gca tct gtt gac cag gct gtt tac gcc atg agc 240
Leu Lys Ala Ala Gly Ala Ser Val Asp Gln Ala Val Tyr Ala Met Ser
65 70 75 80

atc gac gat gtt cgc aat gct tgt gat ctg ttc acc ggc atc ttc gag 288
Ile Asp Asp Val Arg Asn Ala Cys Asp Leu Phe Thr Gly Ile Phe Glu
85 90 95

tcc tcc aac ggc tac gac ggc cgc gtg tcc atc gag gtt gac cca cgt	336
Ser Ser Asn Gly Tyr Asp Gly Arg Val Ser Ile Glu Val Asp Pro Arg	
100 105 110	
atc tct gct gac cgc gac gca acc ctg gct cag gcc aag gag ctg tgg	384
Ile Ser Ala Asp Arg Asp Ala Thr Leu Ala Gln Ala Lys Glu Leu Trp	
115 120 125	
gca aag gtt gat cgt cca aac gtc atg atc aag atc cct gca acc cca	432
Ala Lys Val Asp Arg Pro Asn Val Met Ile Lys Ile Pro Ala Thr Pro	
130 135 140	
ggt tct ttg cca gca atc acc gac gct ttg gct gag ggc atc agc gtt	480
Gly Ser Leu Pro Ala Ile Thr Asp Ala Leu Ala Glu Gly Ile Ser Val	
145 150 155 160	
aac gtc acc ttg atc ttc tcc gtt gct cgc tac cgc gag gtc atc gct	528
Asn Val Thr Leu Ile Phe Ser Val Ala Arg Tyr Arg Glu Val Ile Ala	
165 170 175	
gcg tac atc gag gga atc aag cag gca gct gca aac ggc cac gac gta	576
Ala Tyr Ile Glu Gly Ile Lys Gln Ala Ala Asn Gly His Asp Val	
180 185 190	
tcc aag atc cac tct gtg gct tcc ttc gtc tcc cgc gtc gac gtt	624
Ser Lys Ile His Ser Val Ala Ser Phe Phe Val Ser Arg Val Asp Val	
195 200 205	
gag atc gac aag cgc ctc gag gca atc gga tcc gat gag gct ttg gct	672
Glu Ile Asp Lys Arg Leu Glu Ala Ile Gly Ser Asp Glu Ala Leu Ala	
210 215 220	
ctg cgc ggc aag gca ggc gtt gcc aac gct cag cgc gct tac gct gtg	720
Leu Arg Gly Lys Ala Gly Val Ala Asn Ala Gln Arg Ala Tyr Ala Val	
225 230 235 240	
tac aag gag ctt ttc gac gcc gag ctg cct gaa ggt gcc aac act	768
Tyr Lys Glu Leu Phe Asp Ala Ala Glu Leu Pro Glu Gly Ala Asn Thr	
245 250 255	
cag cgc cca ctg tgg gca tcc acc ggc gtg aag aac cct gcg tac gct	816
Gln Arg Pro Leu Trp Ala Ser Thr Gly Val Lys Asn Pro Ala Tyr Ala	
260 265 270	
gca act ctt tac gtt tcc gag ctg gct ggt cca aac acc gtc aac acc	864
Ala Thr Leu Tyr Val Ser Glu Leu Ala Gly Pro Asn Thr Val Asn Thr	
275 280 285	
atg cca gaa ggc acc atc gac gct gtt ctg gaa ctg ggc aac ctg cac	912
Met Pro Glu Gly Thr Ile Asp Ala Val Leu Glu Leu Gly Asn Leu His	

290	295	300	
ggt gac acc ctg tcc aac tcc gcg gca gaa gct gac gct gtg ttc tcc			960
Gly Asp Thr Leu Ser Asn Ser Ala Ala Glu Ala Asp Ala Val Phe Ser			
305	310	315	320
cag ctt gag gct ctg ggc gtt gac ttg gca gat gtc ttc cag gtc ctg			1008
Gln Leu Glu Ala Leu Gly Val Asp Leu Ala Asp Val Phe Gln Val Leu			
325	330	335	
gag acc gag ggt gtg gac aag ttt gtt gct tct tgg agc gaa ctg ctt			1056
Glu Thr Glu Gly Val Asp Lys Phe Val Ala Ser Trp Ser Glu Leu Leu			
340	345	350	
gag tcc atg gaa gct cgc ctg aag			
Glu Ser Met Glu Ala Arg Leu Lys			
355	360		

<210> 2
 <211> 1080
 <212> DNA
 <213> Corynebacterium glutamicum ATCC31388

<400> 2

atgtctcaca ttgatgatct tgcacagctc ggcacttcca cttggctcga cgacctctcc 60
 cgcgagcgca ttacttcgg caatctcagc caggttattt aggaaaagtc tgttagtcgg 120
 gtcaccacca acccagctat tttcgcagca gcaatgtcca agggcgattc ctacgacgct 180
 cagatcgccag agctcaaggc cgctggcgca tctgttgacc aggctgtta cgccatgagc 240
 atcgacgatg ttgcataatgc ttgtgatctg ttcaccggca tcttcgagtc ctccaaacggc 300
 tacgacggcc gcgtgtccat cgaggttgac ccacgttatct ctgctgaccg cgacgcaacc 360
 ctggctcagg ccaaggagct gtgggcaaag gttgatcgtc caaacgtcat gatcaagatc 420
 cctgcaaccc caggttcttt gccagcaatc accgacgctt tggctgaggg catcagcggt 480
 aacgtcacct tgatcttctc cggtgctcgc taccgcgagg tcacgcgtgc gtacatcgag 540
 ggaatcaagc aggcagctgc aaacggccac gacgtatcca agatccactc tgtggcttcc 600
 ttcttcgtct cccgcgtcga cggtgagatc gacaagcgcc tcgaggcaat cggatccgat 660
 gaggctttgg ctctgcgcgg caaggcaggc gttgccaacg ctcagcgcgc ttacgctgtg 720
 tacaaggagc ttttcgacgc cgccgagctg cctgaagggtg ccaacactca gcgcccactg 780

tggcatcca ccgcgtgaa gaaccctgctacgctgaa ctcttacgt ttccgagctg 840
 gctggtccaa acaccgtcaa caccatgcca gaaggcacca tcgacgctgt tctggaactg 900
 ggcaacctgc acggtgacac cctgtccaa tccgcggcag aagctgacgc tgtttctcc 960
 cagcttggagg ctctggcgt tgacttggca gatgtcttcc aggtcctgga gaccgagggt 1020
 gtggacaagt ttgttgcttc ttggagcgaa ctgcttgagt ccatggaagc tcgcctgaag 1080

<210> 3

<211> 4108

<212> DNA

<213> Corynebacterium glutamicum ATCC31388

<221> CDS

<222> (373)..(2472)

<221> CDS

<222> (2643)..(3722)

<400> 3

tcgagagttt gaaggggtcc gattcgttcc gttcgtgacg ctttgtgagg tttttgacg 60
 ttgcaccgta ttgcttgcg aacattttc tttccttgc gtttttcga gaattttcac 120
 ctacaaaagc ccacgtcaca gctcccagac ttaagattgg tcacacctt gacacatttgc 180
 aaccacagtt gttataaaa tgggttcaac atcactatgg ttagaggtgt tgacgggtca 240
 gattaagcaa agactacttt cgggttagat caccttgcc aaatttgaat caattaacct 300
 aagtcgtaga tctgatcatc ggatctaacg aaaacgaacc aaaacttgg tcccggtta 360
 acccaggaag ga atg acc acc ttg acg ctg tca cct gaa ctt cag gcg ctc 411
 Met Thr Leu Thr Leu Ser Pro Glu Leu Gln Ala Leu
 1 5 10

act gta cgc aat tac ccc tct gat tgg tcc gat gtg gac acc aag gct 459
 Thr Val Arg Asn Tyr Pro Ser Asp Trp Ser Asp Val Asp Thr Lys Ala
 15 20 25

gta gac act gtt gtc ctc gct gca gac gct gta gaa aac tgt ggc 507
 Val Asp Thr Val Arg Val Leu Ala Ala Asp Ala Val Glu Asn Cys Gly
 30 35 40 45

tcc ggc cac cca ggc acc gca atg agc ctg gct ccc ctt gca tac acc 555
 Ser Gly His Pro Gly Thr Ala Met Ser Leu Ala Pro Leu Ala Tyr Thr
 50 55 60

ttg tac cag cgg gtt atg aac gta gat cca cag gac acc aac tgg gca	603
Leu Tyr Gln Arg Val Met Asn Val Asp Pro Gln Asp Thr Asn Trp Ala	
65	70
	75
ggc cgt gac cgc ttc gtt ctt tct tgt ggc cac tcc tct ttg acc cag	651
Gly Arg Asp Arg Phe Val Leu Ser Cys Gly His Ser Ser Leu Thr Gln	
80	85
	90
tac atc cag ctt tac ttg ggt gga ttc ggc ctt gag atg gat gac ctg	699
Tyr Ile Gln Leu Tyr Leu Gly Gly Phe Gly Leu Glu Met Asp Asp Leu	
95	100
	105
aag gct ctg cgc acc tgg gat tcc ttg acc cca gga cac cct gag tac	747
Lys Ala Leu Arg Thr Trp Asp Ser Leu Thr Pro Gly His Pro Glu Tyr	
110	115
	120
	125
cgc cac acc aag ggc gtt gag atc acc act ggc cct ctt ggc cag ggt	795
Arg His Thr Lys Gly Val Glu Ile Thr Thr Gly Pro Leu Gly Gln Gly	
130	135
	140
ctt gca tct gca gtt ggt atg gcc atg gct gct cgt cgt gag cgt ggc	843
Leu Ala Ser Ala Val Gly Met Ala Met Ala Ala Arg Arg Glu Arg Gly	
145	150
	155
cta ttc gac cca acc gct gct gag ggc gaa tcc cca ttc gac cac cac	891
Leu Phe Asp Pro Thr Ala Ala Glu Gly Glu Ser Pro Phe Asp His His	
160	165
	170
atc tac gtc att gct tct gat ggt gac ctg cag gaa ggt gtc acc tct	939
Ile Tyr Val Ile Ala Ser Asp Gly Asp Leu Gln Glu Gly Val Thr Ser	
175	180
	185
gag gca tcc tcc atc gct ggc acc cag cag ctg ggc aac ctc atc gtg	987
Glu Ala Ser Ser Ile Ala Gly Thr Gln Gln Leu Gly Asn Leu Ile Val	
190	195
	200
	205
ttc tgg gat gac aac cgc atc tcc atc gaa gac aac act gag atc gct	1035
Phe Trp Asp Asp Asn Arg Ile Ser Ile Glu Asp Asn Thr Glu Ile Ala	
210	215
	220
ttc aac gag gac gtt gtt gct cgt tac aag gct tac ggc tgg cag acc	1083
Phe Asn Glu Asp Val Val Ala Arg Tyr Lys Ala Tyr Gly Trp Gln Thr	
225	230
	235
att gag gtt gag gct ggc gag gac gtt gca gca atc gaa gct gca gtg	1131
Ile Glu Val Glu Ala Gly Glu Asp Val Ala Ala Ile Glu Ala Ala Val	
240	245
	250
gct gag gct aag aag gac acc aag cga cct acc ttc atc cgc gtt cgc	1179
Ala Glu Ala Lys Lys Asp Thr Lys Arg Pro Thr Phe Ile Arg Val Arg	

255	260	265	
acc atc atc ggc ttc cca gct cca acc atg atg aac acc acc ggt gct gtg Thr Ile Ile Gly Phe Pro Ala Pro Thr Met Met Asn Thr Gly Ala Val			1227
270	275	280	285
cac ggt gct gct ctt ggc gca gct gag gtt gca gca acc aag act gag His Gly Ala Ala Leu Gly Ala Ala Glu Val Ala Ala Thr Lys Thr Glu			1275
290	295	300	
ctt gga ttc gat cct gag gct cac ttc gcg atc gac gat gag gtt atc Leu Gly Phe Asp Pro Glu Ala His Phe Ala Ile Asp Asp Glu Val Ile			1323
305	310	315	
gct cac acc cgc tcc ctc gca gag cgc gct gca cag aag aag gct gca Ala His Thr Arg Ser Leu Ala Glu Arg Ala Ala Gln Lys Lys Ala Ala			1371
320	325	330	
tgg cag gtc aag ttc gat gag tgg gca gct gcc aac cct gag aac aag Trp Gln Val Lys Phe Asp Glu Trp Ala Ala Asn Pro Glu Asn Lys			1419
335	340	345	
gct ctg ttc gat cgc ctg aac tcc cgt gag ctt cca gcg ggc tac gct Ala Leu Phe Asp Arg Leu Asn Ser Arg Glu Leu Pro Ala Gly Tyr Ala			1467
350	355	360	365
gac gag ctc cca aca tgg gat gca gat gag aag ggc gtc gca act cgt Asp Glu Leu Pro Thr Trp Asp Ala Asp Glu Lys Gly Val Ala Thr Arg			1515
370	375	380	
aag gct tcc gag gct gca ctt cag gca ctg ggc aag acc ctt cct gag Lys Ala Ser Glu Ala Ala Leu Gln Ala Leu Gly Lys Thr Leu Pro Glu			1563
385	390	395	
ctg tgg ggc ggt tcc gct gac ctc gca ggt tcc aac aac acc gtg atc Leu Trp Gly Gly Ser Ala Asp Leu Ala Gly Ser Asn Asn Thr Val Ile			1611
400	405	410	
aag ggc tcc cct tcc ttc ggc cct gag tcc atc tcc acc gag acc tgg Lys Gly Ser Pro Ser Phe Gly Pro Glu Ser Ile Ser Thr Glu Thr Trp			1659
415	420	425	
tct gct gag cct tac ggc cgt aac ctg cac ttc ggt atc cgt gag cac Ser Ala Glu Pro Tyr Gly Arg Asn Leu His Phe Gly Ile Arg Glu His			1707
430	435	440	445
gct atg gga tcc atc ctc aac ggc att tcc ctc cac ggt ggc acc cgc Ala Met Gly Ser Ile Leu Asn Gly Ile Ser Leu His Gly Gly Thr Arg			1755
450	455	460	

cca tac ggt gga acc ttc ctc atc ttc tcc gac tac atg cgt cct gca	1803
Pro Tyr Gly Gly Thr Phe Leu Ile Phe Ser Asp Tyr Met Arg Pro Ala	
465 470 475	
gtt cgt ctt gca gct ctc atg gag acc gac gct tac tac gtc tgg acc	1851
Val Arg Leu Ala Ala Leu Met Glu Thr Asp Ala Tyr Tyr Val Trp Thr	
480 485 490	
cac gac tcc atc ggt ctg ggc gaa gat ggc cca acc cac cag cct gtt	1899
His Asp Ser Ile Gly Leu Gly Glu Asp Gly Pro Thr His Gln Pro Val	
495 500 505	
gaa acc ttg gct gcg ctg cgc gcc atc cca ggt ctg tcc gtc ctg cgt	1947
Glu Thr Leu Ala Ala Leu Arg Ala Ile Pro Gly Leu Ser Val Leu Arg	
510 515 520 525	
cct gca gat gcg aat gag acc gcc cag gct tgg gct gca gca ctt gag	1995
Pro Ala Asp Ala Asn Glu Thr Ala Gln Ala Trp Ala Ala Ala Leu Glu	
530 535 540	
tac aag gaa ggc cct aag ggt ctt gca ctg acc cgc cag aac gtt cct	2043
Tyr Lys Glu Gly Pro Lys Gly Leu Ala Leu Thr Arg Gln Asn Val Pro	
545 550 555	
gtt ctg gaa ggc acc aag gag aag gct gct gaa ggc gtt cgc cgc ggt	2091
Val Leu Glu Gly Thr Lys Glu Ala Ala Glu Gly Val Arg Arg Gly	
560 565 570	
ggc tac gtc ctg gtt gag ggt tcc aag gaa acc cca gat gtg atc ctc	2139
Gly Tyr Val Leu Val Glu Gly Ser Lys Glu Thr Pro Asp Val Ile Leu	
575 580 585	
atg ggc tcc ggc tcc gag gtt cag ctt gca gtt aac gct gcg aaa gct	2187
Met Gly Ser Gly Ser Glu Val Gln Leu Ala Val Asn Ala Ala Lys Ala	
590 595 600 605	
ctg gaa gct gag ggc gtt gca gct cgc gtt tca gtt cct tgc atg	2235
Leu Glu Ala Glu Gly Val Ala Ala Arg Val Val Ser Val Pro Cys Met	
610 615 620	
gat tgg ttc cag gag cag gac gca gag tac atc gag tcc gtt ctg cct	2283
Asp Trp Phe Gln Glu Gln Asp Ala Glu Tyr Ile Glu Ser Val Leu Pro	
625 630 635	
gca gct gtg acc gct cgt gtg tct gtt gaa gct ggc atc gca atg cct	2331
Ala Ala Val Thr Ala Arg Val Ser Val Glu Ala Gly Ile Ala Met Pro	
640 645 650	
tgg tac cgc ttc ttg ggc acc cag ggc cgt gct gtc tcc ctt gag cac	2379
Trp Tyr Arg Phe Leu Gly Thr Gln Gly Arg Ala Val Ser Leu Glu His	

655	660	665	
ttc ggt gct tct gcg gat tac cag acc ctg ttt gag aag ttc ggc atc Phe Gly Ala Ser Ala Asp Tyr Gln Thr Leu Phe Glu Lys Phe Gly Ile 670 675 680 685			2427
acc acc gat gca gtc gtg gca gcg gcc aag gac tcc att aac agt Thr Thr Asp Ala Val Val Ala Ala Lys Asp Ser Ile Asn Ser 690 695 700			2472
taattgccct gctgtttta gcttcaaccc gggcagttat gattctccgg aattttatttg 2532			
ccccgggttg ttgttgttaa tcggtacaaa gggtcttaag cacatccctt acttgcctgc 2592			
tctcctttag cacagttcaa gaacaattct tttaaggaaa atttagttc atg tct His Ile Asp Asp Leu Ala Gln Leu Gly Thr Ser Thr Trp Leu Asp Asp 5 10 15			2648
Met Ser 1			
cac att gat gat ctt gca cag ctc ggc act tcc act tgg ctc gac gac Leu Ser Arg Glu Arg Ile Thr Ser Gly Asn Leu Ser Gln Val Ile Glu 20 25 30			2696
ctc tcc cgc gag cgc att act tcc ggc aat ctc agc cag gtt att gag Leu Ser Arg Glu Arg Ile Thr Ser Gly Asn Leu Ser Gln Val Ile Glu 20 25 30			2744
gaa aag tct gta gtc ggt gtc acc acc aac cca gct att ttc gca gca Glu Lys Ser Val Val Gly Val Thr Thr Asn Pro Ala Ile Phe Ala Ala 35 40 45 50			2792
gca atg tcc aag ggc gat tcc tac gac gct cag atc gca gag ctc aag Ala Met Ser Lys Gly Asp Ser Tyr Asp Ala Gln Ile Ala Glu Leu Lys 55 60 65			2840
gcc gct ggc gca tct gtt gac cag gct gtt tac gcc atg agc atc gac Ala Ala Gly Ala Ser Val Asp Gln Ala Val Tyr Ala Met Ser Ile Asp 70 75 80			2888
gat gtt cgc aat gct tgt gat ctg ttc acc ggc atc ttc gag tcc tcc Asp Val Arg Asn Ala Cys Asp Leu Phe Thr Gly Ile Phe Glu Ser Ser 85 90 95			2936
aac ggc tac gac ggc cgc gtg tcc atc gag gtt gac cca cgt atc tct Asn Gly Tyr Asp Gly Arg Val Ser Ile Glu Val Asp Pro Arg Ile Ser 100 105 110			2984
gct gac cgc gac gca acc ctg gct cag gcc aag gag ctg tgg gca aag Ala Asp Arg Asp Ala Thr Leu Ala Gln Ala Lys Glu Leu Trp Ala Lys 115 120 125 130			3032
gtt gat cgt cca aac gtc atg atc aag atc cct gca acc cca ggt tct 3080			

Val Asp Arg Pro Asn Val Met Ile Lys Ile Pro Ala Thr Pro Gly Ser			
135	140	145	
ttg cca gca atc acc gac gct ttg gct gag ggc atc agc gtt aac gtc.	3128		
Leu Pro Ala Ile Thr Asp Ala Leu Ala Glu Gly Ile Ser Val Asn Val			
150	155	160	
acc ttg atc ttc tcc gtt gct cgc tac cgc gag gtc atc gct gcg tac	3176		
Thr Leu Ile Phe Ser Val Ala Arg Tyr Arg Glu Val Ile Ala Ala Tyr			
165	170	175	
atc gag gga atc aag cag gca gct gca aac ggc cac gac gta tcc aag	3224		
Ile Glu Gly Ile Lys Gln Ala Ala Asn Gly His Asp Val Ser Lys			
180	185	190	
atc cac tct gtg gct tcc ttc gtc tcc cgc gtc gac gtt gag atc	3272		
Ile His Ser Val Ala Ser Phe Val Ser Arg Val Asp Val Glu Ile			
195	200	205	210
gac aag cgc ctc gag gca atc gga tcc gat gag gct ttg gct ctg cgc	3320		
Asp Lys Arg Leu Glu Ala Ile Gly Ser Asp Glu Ala Leu Ala Leu Arg			
215	220	225	
ggc aag gca ggc gtt gcc aac gct cag cgc gct tac gct gtg tac aag	3368		
Gly Lys Ala Gly Val Ala Asn Ala Gln Arg Ala Tyr Ala Val Tyr Lys			
230	235	240	
gag ctt ttc gac gcc gag ctg cct gaa ggt gcc aac act cag cgc	3416		
Glu Leu Phe Asp Ala Ala Glu Leu Pro Glu Gly Ala Asn Thr Gln Arg			
245	250	255	
cca ctg tgg gca tcc acc ggc gtg aag aac cct gcg tac gct gca act	3464		
Pro Leu Trp Ala Ser Thr Gly Val Lys Asn Pro Ala Tyr Ala Ala Thr			
260	265	270	
ctt tac gtt tcc gag ctg gct ggt cca aac acc gtc aac acc atg cca	3512		
Leu Tyr Val Ser Glu Leu Ala Gly Pro Asn Thr Val Asn Thr Met Pro			
275	280	285	290
gaa ggc acc atc gac gct gtt ctg gaa ctg ggc aac ctg cac ggt gac	3560		
Glu Gly Thr Ile Asp Ala Val Leu Glu Leu Gly Asn Leu His Gly Asp			
295	300	305	
acc ctg tcc aac tcc gcg gca gaa gct gac gct gtg ttc tcc cag ctt	3608		
Thr Leu Ser Asn Ser Ala Ala Glu Ala Asp Ala Val Phe Ser Gln Leu			
310	315	320	
gag gct ctg ggc gtt gac ttg gca gat gtc ttc cag gtc ctg gag acc	3656		
Glu Ala Leu Gly Val Asp Leu Ala Asp Val Phe Gln Val Leu Glu Thr			
325	330	335	

gag ggt gtg gac aag ttt gtt gct tct tgg agc gaa ctg ctt gag tcc 3704
Glu Gly Val Asp Lys Phe Val Ala Ser Trp Ser Glu Leu Leu Glu Ser
340 345 350

atg gaa gct cgc ctg aag tagaatcagc acgctgcattc agtaacggcg 3752
Met Glu Ala Arg Leu Lys
355 360

acatgaaaatc gaatttagttc gatcttatgt ggccgttaca catctttcat taaagaaaagg 3812
atcggtgacgc taccatcggt agcacaaaca cgaccccttc cagctggaca aacccactgc 3872
gcgacccgca ggataaacga ctccccgca tcgctggccc ttccggcatg gtgatcttcg 3932
gtgtcactgg cgacttgct cgaaggaagc tgctccccgc catttatgtat ctagcaaacc 3992
gcggattgct gcccccagga ttctcggttgg tagttacgg ccggccgcgaa tggtccaaag 4052
aagactttga aaaatacgta cgcgatgccg caagtgttgg tgctcgtaacg gaattc 4108